

UC Irvine

ICS Technical Reports

Title

Survey of traffic control schemes and error control schemes for ATM networks

Permalink

<https://escholarship.org/uc/item/79t9k7g4>

Authors

Bae, Jaime Jungok
Suda, Tatsuya

Publication Date

1991

Peer reviewed

Notice: This Material
may be protected
by Copyright Law
(Title 17 U.S.C.)

Z
699
C3
no. 91-46

**Survey of Traffic Control Schemes and
Error Control Schemes for ATM Networks***

Jaime Jungok Bae and Tatsuya Suda

Technical Report No. 91-46

Department of Information and Computer Science
University of California, Irvine
Irvine, CA 92717
(phone) 714-856-5474
(fax) 714-856-4056

* This material is based upon work supported by the National Science Foundation under Grant No. NCR-8907909. This research is also in part supported by University of California MICRO program.

Abstract

Among the techniques proposed for B-ISDN transfer mode, ATM concept is considered to be the most promising transfer technique because of its flexibility and efficiency. This paper surveys and reviews a number of topics related to ATM networks. Those topics cover congestion control, provision of multiple classes of traffic, and error control. Due to the nature of ATM networks, those issues are far more challenging than in conventional networks. Some of the more promising solutions to those issues are surveyed, and the corresponding results on performance are summarized. Future research problems in ATM protocol aspect are also presented.

1. Introduction

Due to the recent technological advances, especially in fiber optics and micro electronics, a great deal of attention has been paid to the Broadband ISDN (B-ISDN). The Broadband ISDN provides much higher transmission rates than the conventional Narrowband ISDN does. Such a network should be able to support a wide variety of traffics requiring bandwidth ranging from a few Killo-bits per second (e.g., a slow terminal) to several hundred Mega-bits per second (e.g., moving image data). Some input traffic (e.g., interactive data and spoken voice) are highly bursty, and some (e.g., file-transfers) are continuous. Therefore, such a network needs to satisfy varying input traffics. The Broadband ISDN should also be able to satisfy diverse service and performance requirements. For instance, real-time voice requires rapid transfer through a network, but loss of small amount of voice information is tolerable. In many data applications, however, real-time delivery is not of primary importance, but high throughput and strict error control are required. Some services, such as video conferencing communications, require error-free transmission as well as rapid transfer [1]. Therefore, the Broadband ISDN should be able to facilitate a wide range of source bit-rates and various types of services with different traffic characteristics such as voice, data and video. The Broadband ISDN should also be able to support unpredictable future services. Some of the examples of possible types of future services are [2,3]:

- video surveillance
- high-quality broadband video telephony,
- high-quality broadband video conferencing,
- video/document retrieval service,
- high-speed unrestricted digital information transmission,
- TV distribution with existing standards quality and high-definition quality, and
- broadband videotex.

Therefore, such a network should be very flexible so that it can be adapted to all expected and unexpected changes in the future.

Several techniques have been proposed for the switching and multiplexing schemes ("transfer mode") for B-ISDN. They are the Synchronous Transfer Mode (STM) based on circuit-switching, the Asynchronous Transfer Mode (ATM) based on packet-switching, and hybrid

of ATM and STM. Among them, the ATM is considered to be the most promising transfer technique, and thus this paper surveys and reviews a number of topics related to ATM networks.

The STM was initially assumed to be the appropriate transfer mode for B-ISDN by many because of its compatibility with existing systems. In STM, the transmission medium has its bandwidth organized in a periodic frame, and an STM channel is identified by the position of its time slots within a synchronous structure (see Figure 1(a)). In other words, as in circuit switching, a channel is given a fixed slot location within a frame for the duration of a call. Therefore, STM is suitable for fixed-rate services. However, B-ISDN needs to support a variety of services including bursty services. Since in STM, channel bandwidths (slots) are allocated based on the peak transfer rates, bandwidths are wasted during the period no information is transported, and thus, STM techniques cannot transport these services efficiently.

ATM attempts to eliminate this inflexibility of STM. ATM combines the advantages of circuit switching and packet switching. In ATM, information flow is organized into fixed-size blocks, called "cells", the cell transmission time being equal to a slot length. Specific time slots are not assigned to channels (see Figure 1(b)), and instead, slots are allocated to services on demand. Each cell consists of a header and an information field. These cells are transmitted over a virtual circuit (i.e., connection-oriented), and routing is performed based on a label contained in the header. This label is comparable to the slot position in STM for channel identification. The fundamental difference to STM is that the time slots are used in an asynchronous (on demand) manner by the various logical channels. In ATM, unlike in the STM, no bandwidth is consumed unless information is actually being transported.

Hybrid solutions are also proposed for the transfer mode for B-ISDN. There are two types of hybrids - horizontal hybrid and vertical hybrid. In a vertical hybrid, STM and ATM techniques are employed at different levels of an interface, whereas in a horizontal hybrid, both techniques coexist at the same level. An example of a vertical hybrid is an interface that carries ATM cells within a frame structure which is divided into two areas using STM techniques. An example of a horizontal hybrid is an interface that can support services on both STM and ATM channels. In this case, some channels are identified by time slot, others by labels [62]. The practical utility of horizontal hybrid has not been investigated.

Among the techniques proposed for B-ISDN transfer mode, ATM concept is considered to be the most promising transfer technique because of its flexibility and efficiency. ATM is flexible since ATM does not use fixed-rate connections, and thus variable bit rate services can be accommodated easily. Because of the dynamic allocation of bandwidth, the ATM is also sensitive to variations in demand for services. ATM can also gain bandwidth efficiency by statistically multiplexing bursty traffic generated from several sources and service types. Each service may not require continuous allocation of its peak bandwidth, allowing more services to statistically share the resources. ATM can also guarantee acceptable performance for continuous-bit-rate services. Operating in a deterministic mode (i.e., allocating bandwidth based on the peak rates), ATM can support real-time continuous-bit-rate services, given that sufficient resources are guaranteed. For a specific continuous-bit-rate service, ATM may not be as efficient as STM in terms of bandwidth utilization because of additional overhead

contained in the header of each cell, however, it may be offset by its overall merits.

The organization of this paper is as follows. Congestion control, provision of multiple classes of traffic, and error control are discussed in sections 2, 3, and 4 respectively. Finally, in section 5, a brief conclusion and possible future research problems are presented.

2. Congestion Control

In an ATM network, most traffic sources have bursty nature. A bursty source may generate cells at a near-peak rate for a short period of time, and a second later, it may become idle contributing no traffic on the network. Statistical multiplexing can be used to gain bandwidth efficiency. In this environment, if all the calls become active simultaneously, network becomes congested. Congestion can also happen due to unforeseen traffic characteristics and network component failures.

Congestion control is a challenge in an ATM network due to the effects of high speed channels. High speed channels drastically limit the congestion control schemes applicable to ATM networks. As an example, consider two adjacent switching nodes, A and B, linked by a 100 Km cable. Assume 1000 bit long packets and a typical propagation delay time of 5 μ sec per 1 Km of a cable. Consider the following scenario. Assume a 1 Mbits/sec channel speed. One packet transmission time, then, becomes $(1000 \text{ bits})/(1 \text{ Mbits/sec}) = 1 \text{ msec}$. Node A starts transmitting a packet. It takes 500 μ sec for the electric signal to propagate to node B. Thus, when the first bit of the packet reaches B, A is transmitting the 500-th bit of the same packet. Let's replace the channel with a 1 Gbits/sec fiber optic cable. Then, packet transmission time reduces to $(1000 \text{ bits})/(1 \text{ Gbits}) = 1 \mu$ sec, while the propagation delay time remains the same. Again, A starts transmitting a packet. This time, when the first bit of the packet arrives at B, A is transmitting the 500-th packet (not the 500-th bit of the first packet). 500 packets are already on the channel propagating towards B. This example suggests that control schemes to adjust A's input rate based on feedback from B may not work in ATM networks.

High speed channels also increase the ratio of processing time to transmission time. For instance, if a 1 Mbits/sec channel speed is used for 1000 bit long packets, one packet transmission time is 1 msec, whereas if a 1 Gbits/sec channel speed is used instead, one packet transmission time becomes 1 μ sec. Therefore, the network bottleneck is moved from the transmission links to the switching nodes. To match the rapid channel speed, in ATM networks, communication protocols are simplified, reducing protocol processing time. Most of the link-to-link layer protocols have been simplified or removed, and pushed to higher edge-to-edge layers. This makes it difficult to implement link-by-link based congestion control schemes.

Due to the above reasons, congestion schemes developed for existing networks may not work. In existing X.25 packet networks, congestion control is performed by a window flow control mechanism. Window flow control mechanisms keep track of the number of packets pending in the network for each virtual circuit using packet-by-packet acknowledgement scheme. Even though this mechanism can smooth resource utilization, it is difficult for the network to satisfy user throughput requirements. This conventional flow control mechanism is not appropriate in an ATM environment for the following reasons [63]:

- (1) Some voice or video communications require guaranteed user throughput.
- (2) It is difficult to use window-type flow control for the high bit rate ATM network since high bit rate services require enormous windows and processing burden including time-out recovery.

Therefore, a new concept is required for the congestion control in an ATM environment.

We define a congestion control as a control mechanism to prevent the congestion as well as to relieve the congestion after it happens. Note that this definition differs from the conventionally used definition, which only includes control mechanism to relieve the congestion after it happens. Various congestion control approaches have been proposed, and they can be divided into two categories, reactive control and preventive control. Reactive control reacts to the congestion when it happens. The preventive control tries to prevent congestion before it happens. The objective of preventive controls is to ensure a priori that the network traffic intensity will normally never reach the level to cause unacceptable congestion. These two approaches will be discussed in detail in the following subsections 2.1 and 2.2.

2.1. Reactive Control

At the onset of congestion, reactive control instructs the source nodes to throttle their traffic flow by giving feedback to the source nodes. There are several problems with the reactive control in high-speed networks [20]:

- Since high capacity links suffer a sudden degradation in performance with increasing load, when the congestion is detected, it may be too late to react effectively. Even if this were possible, a bursty traffic may cause the rapid buffer level fluctuations which could unnecessarily trigger controls.
- As discussed earlier, effects of high speed channels make the propagation delay of approximately $5 \mu\text{s}/\text{km}$ required to send the choke packet to the source significant. Therefore, the feedback information cannot be used in effective way.
- Different types of services may need to be treated differently. However, this approach of throttling all traffic sources routed through the congested point regardless of the type of services fails to take this into account.

A possible solution to overcome the difficulty caused by slow feedback is to perform a reactive control only at the edge of the network as in [19]. In this case, the propagation delay will not be a problem since the distance feedback message propagates is short.

However, because of the problems discussed above, reactive control is, in general, not effective in an ATM network, and thus, very little work has been addressed to the reactive control. The preventive control which will be discussed in the next subsection may be more effective in ATM.

2.2. Preventive Control

Unlike in reactive control where control is invoked after congestion happens, the preventive control does not wait until the congestion is observed, but rather anticipates the occurrence of congestion and prevents the network from reaching the unacceptable conges-

tion level. The common approach to preventive controls is to control the traffic flow admitted into the network at the access nodes. This approach is more effective in connection-oriented transport than in connectionless since a decision to admit a connection's traffic can be made based on knowledge of the state of the route which the traffic would follow [20]. An ATM is connection-oriented (i.e., packets belonging to the same connection take the same route throughout the network), and in ATM networks, an admission procedure is invoked before the connection set-up.

With connection-oriented transport, the preventive control can be performed in two ways - admission control and bandwidth enforcement. The admission control determines whether to accept or reject a new call at the time of call set-up based on the user's traffic descriptions and current network load so as to satisfy performance requirements. The bandwidth enforcement polices individual connections to ensure the actual traffic flow conforms with that specified at call establishment. In the next subsections, subsections 2.2.1 and 2.2.2, admission control and bandwidth enforcement are further discussed respectively.

Holtzman has proposed new and very different approaches to a preventive control in [21]. His approaches deal with how to cope with traffic uncertainties. His work is very interesting in the sense that it is the first work dealing with this subject, and the approaches taken in his work are new. Thus, his approaches are included in section 2.2.3.

2.2.1. Admission Control

Admission control decides whether to accept or reject a new connection based on whether the required performance can be maintained. When a new connection is requested, the network examines its service requirements (e.g, acceptable cell transmission delay and loss probability) and traffic characteristics (e.g., peak rate, average rate, etc.). The network then examines the current load and decides whether or not to accept the new connection.

Three major research issues in admission control are:

- What traffic parameters (traffic descriptors) are required to accurately predict network performance?
- What criteria should the network use to decide whether or not to accept a new connection?
- How does network performance depend on various traffic parameters?

In the following, these three issues are discussed.

Traffic Descriptors

When a new connection is requested, the network needs to know the traffic characteristics of the new connection in order to accurately predict its ability to maintain a certain performance level. A set of traffic descriptors given from a user to a network should include sufficient parameters so that the network can accurately determine the user's traffic characteristics. However, for simplicity's sake a set of traffic descriptors should include the fewest possible parameters.

The peak bit rate, the average bit rate, and a measure of burstiness are the most commonly used parameters for traffic descriptors. Among them, "burstiness" is the most impor-

tant parameter, especially in an ATM network where most traffic sources are highly bursty. Burstiness is a parameter which describes how densely or sparsely cell arrivals occur. It is well known that burstiness plays a critical role in determining network performance ; however, consensus is yet to be reached concerning an appropriate way to describe the burstiness of a traffic source. Possible definitions of burstiness proposed include:

1. The ratio of peak bit rate to average bit rate [22 - 25]
2. The average burst length, i.e., the mean duration of the time interval during which the traffic source transmits at the peak rate [26]
3. Burst factor defined as the average number of bits accumulated in a buffer during a burst, namely, (peak bit rate - average service bit rate) \times average burst length [27]
4. Cell jitter ratio defined as the variance-to-mean ratio of the cell interarrival times, namely, $\text{Var}[\text{cell interarrival times}] / E[\text{cell interarrival times}]$ [28]
5. The squared coefficient of variation of the interarrival times, namely, $\text{Var}[\text{cell interarrival times}] / E^2[\text{cell interarrival times}]$ [29]
6. Peakedness defined as the variance-to-mean ratio of the number of busy servers in a fictitious infinite server group [30]

Deciding the best way to describe the burstiness is a very difficult task which needs to be studied further. The authors of this paper believe that the burst length should somehow be taken into account since it significantly affects the performance. In [5, 22, 26, 27, 31], it is shown that the longer the burst length, the worse the network performance becomes; namely, the cell loss probability becomes larger and the cell transmission delay becomes longer. The effect of the average burst length is also examined in [25]. It is shown that with longer bursts, statistical multiplexing becomes less effective, and thus fewer active sources can be supported for a given amount of bandwidth. The authors also believe that more than one parameter may be necessary to describe burstiness.

In [32], a new traffic descriptor is proposed. In this paper, the difficulty of using the peak bit rate or the average bit rate as a traffic descriptor uniformly across the different types of traffic is realized. If the peak bit rate is used regardless of the type of traffic, a large portion of bandwidth will be wasted, especially when the network traffic is bursty. On the other hand, if the average bit rate is used regardless of the type of traffic, the continuous-bit-oriented (CBO) traffic will suffer severe performance degradation. In [32], a new bit rate, called the effective bit rate, is proposed. An effective bit rate is defined as a fraction of the peak bit rate, namely, effective bit rate = (peak bit rate) $\times a$, where a is a constant. The value of a is determined based on the traffic characteristics of the source. By changing the value of a , we can improve the network resource utilization. Further study is required to determine an appropriate value of a for different types of traffic.

Decision Criteria

The cell transmission delays and the cell loss probabilities, because they are good indications of the degree of network congestion, are the most commonly used decision criteria in admission control. When transmission delays and cell loss probabilities are applied in admission control, their long-term-time-averaged values have been used in the past [22, 24,

26, 31 - 33]. Using a long-term-time-averaged value, however, may not be sufficient in an ATM network because here the network traffic can change rapidly and dynamically, forcing the network to move from one degree of congestion to another. Figure 2 [4] sketches how the cell loss probability changes in an ATM network as a function of time. In this figure, the number of active calls jumps from a at time t_0 , to b at time t_1 , and to c at time t_2 . At time t_3 , the number of active calls decreases to b . The solid curve in the figure shows the time-dependent behavior of the cell loss probability. For instance, when the number of active calls increases to b at time t_1 , the network responds to the change and starts losing a large number of cells; gradually, the network goes up to the next level of congestion and reaches the value of the cell loss probability in steady state $P_{loss}(b)$. When another increase occurs at time t_2 , the network responds again, gradually reaching the steady state, and so on. When the network traffic is highly bursty and changes dynamically, temporal network congestion can occur, and it is possible that a large number of cells are lost during congestion periods, even when the long-term-time-averaged value of loss rate is kept small. In voice communication, for example, this burst loss of voice cells may cause noticeable performance degradation (clicks) at a destination user. Therefore, some decision criteria which take the temporal behavior of the network into account may be needed.

In [4], an instantaneous cell loss probability is proposed and used as a decision criterion to consider the temporal behavior of a network. An instantaneous cell loss probability is a time-dependent cell loss probability (function of slot position or time), not the value averaged over a long period of time. The solid curve in Figure 2 shows the instantaneous cell loss probability. In [4], the instantaneous cell loss probability is approximated by its steady state value (dashed lines in Figure 2), and an approximate analysis is developed. A new connection is accepted by the network only when the instantaneous cell loss rate is kept below a threshold value at each switching node for longer than a predetermined percentage of time.

In [4], the ineffectiveness of using the long-term-time averaged cell loss probability as a decision criterion is demonstrated through numerical examples using realistic parameter values. It is shown that network congestion can last for a length of time on the order of a hundred milliseconds even when the long-term-time-averaged cell loss probability is kept small. In voice conversation, this congestion period is comparable to a burst (talkspurt) length, and thus a whole talkspurt can be lost during congestion. It is also shown that this burst cell loss can be avoided by using the instantaneous cell loss probability as a decision criterion in admission control.

In [6], the insufficiency of measuring only the long-term-time averaged cell loss probability is discussed further, and the temporal behavior of voice cell loss probability is studied. Under the realistic parameter values, it is found that the cell loss rate changes slowly and remains at zero most of the time. However, once congestion occurs and the cell loss probability becomes large, the cell loss probability may remain large for a long period, causing voice distortion perceptible at the receiver. It is shown that the average cell loss probability within a blocking period (i.e., the time period during which the buffer is full, and thus cells are blocked) is much larger than the long-term-time averaged cell loss rate. Therefore, the long-term-time averaged cell loss probability does not reflect the temporal behavior of voice cell loss, and it is not sufficient to measure voice distortion incurred.

Effects of Traffic parameters on ATM Network Performance

One of the important research issues in admission control is to investigate the effect of various traffic parameters on network performance.

In [4, 22, 26, 27, 31], the effects of statistical multiplexing of bursty sources in an ATM network are investigated. They investigate how the performance (the cell loss probability and the average delay time) varies as a function of various parameters, such as the number of sources, the peak bit rate, and the burstiness of the sources. Some of the common observations made in these papers follow:

- The average burst length is a very important parameter. As the average burst length increases, the performance degrades, i.e., the cell loss probability and delay time increase significantly [22, 26, 27, 31].
- As the peak rate of each source is increased, the cell loss probability increases [26, 27]. This should be intuitively clear.
- In the case where homogeneous sources are multiplexed, if the offered load (i.e., the number of sources \times mean bit rate of each source) is kept constant, the cell loss probability decreases as the number of sources multiplexed increases. The reason for this is that when the number of sources multiplexed increases (keeping the offered load constant), the mean bit rate of each source decreases. The mean bit rate is a product of peak bit rate and the fraction of time in which a source is in the active-state (i.e., the state in which a source is transmitting at the peak rate). Therefore, the reduction in the mean bit rate means the reduction in either the peak bit rate or the burst length (or both). In either case, the cell loss probability decreases [22, 27].
- In the case where heterogeneous sources are multiplexed, high-bit-rate sources dominate the performance; an increase in high-bit-rate traffic causes more significant increases in the cell loss probability than does an increase in low-bit-rate traffic [31]. A similar observation is made in the case where homogeneous sources are multiplexed; when high-bit-rate sources are multiplexed, the fluctuation in the cell loss is larger than when low-bit-rate sources are multiplexed [4]. This is due to the fact that because of the high bit rate, the number of traffic sources which can be multiplexed on one link is rather limited and not large enough to smooth out the bursty nature of each call.
- The cell loss probability decreases as the offered load decreases [22, 31]. Thus, a very efficient way to lower the cell loss probability is to decrease the offered load by providing larger bandwidth. This is only possible, however, if one can assume that bandwidth is negligibly cheap.

In [25], the effects of traffic parameters on the network performance are investigated and a method is proposed to calculate the bandwidth required to satisfy a given performance requirement. Two different cases are considered: the case where homogeneous traffic sources are multiplexed and the case where heterogeneous sources are multiplexed. In both cases, the peak bit rate (B_p), a measure of burstiness (b) defined as the peak-to-mean bit rate ratio ($\frac{B_p}{B_m}$, where B_m is the mean bit rate), and the mean number of cells (L) generated from a burst are used as traffic descriptors. In the following, we summarize the bandwidth assignment rule proposed for the homogeneous traffic case.

In the case where homogeneous traffic sources are multiplexed, the bandwidth required to satisfy a given cell loss requirement is calculated by

$$W = n \frac{Bp}{b} R(b, n, L) \quad (7)$$

where n is the number of active traffic sources; $n \frac{Bp}{b} (= nB_m)$ is the offered traffic; and $R(b, n, L)$ is a coefficient whose value depends on the triplet (b, n, L) . $R(b, n, L)$ is called an expansion factor, and its value is obtained by performing a single simulation for each triplet (b, n, L) for a given cell loss requirement. In this paper, a cell loss requirement of 10^{-5} is assumed. This cell loss probability is rather large to be used in a real system; this cell loss probability is used because the simulation run time prohibits the choice of a more realistic cell loss probability of 10^{-9} .

Using eq.(7), if the offered traffic ($n \frac{Bp}{b}$) and the expansion factor $R(b, n, L)$ are given, required bandwidth W can be determined. The remaining question is whether the expansion factor is a function of a triplet (b, n, L) , or a function of a quadruplet (Bp, Bm, n, L) . In [25], it is claimed that a triplet (b, n, L) is sufficient to determine the expansion factor, and it is supported by examining simulation results for two cases: $Bp = 10$ Mbits/sec or 2 Mbits/sec. To determine that this approach is truly valid, more cases should be examined.

The approach proposed in this paper considers the burstiness of the traffic and uses the peak-to-mean bit rate ratio ($b = \frac{Bp}{Bm}$) and the mean number of cells generated in a burst (L) to determine required bandwidth (W). Even though the approach of using $R(b, n, L)$ to calculate required bandwidth is simpler than the approach in which the quadruplet (Bp, Bm, n, L) is used, it has the following problem. To implement this approach, the values of $R(b, n, L)$ need to be precomputed through the simulation and stored in each node. Therefore, the number of possible combinations of (b, n, L) needs to be tractably small. This may limit the size of the network to which this approach can apply.

2.2.2. Bandwidth Enforcement¹

Since users may deliberately exceed the traffic volume declared at the call set up (i.e., values of their traffic descriptors), and thus easily overload the network, admission control alone is not sufficient. After a connection is accepted, traffic flow of the connection must be monitored to insure that the actual traffic flow conforms with that specified at call establishment. For this purpose, the bandwidth enforcement mechanism is implemented at the edges of the network. Once a violation is detected, the traffic flow is enforced by discarding and/or buffering violating cells.

A Leaky Bucket method [27, 34 - 36] is one of the typical bandwidth enforcement mechanisms used for ATM networks; this method can enforce the average bandwidth and the burst factor of a traffic source. One possible implementation of a Leaky Bucket method is to control the traffic flow by means of tokens. A queueing model for the Leaky Bucket method is illustrated in Figure 3 [37]. An arriving cell first enters a queue. If the queue is full, cells are simply discarded. To enter the network, a cell must first obtain a token from a token-pool; if there is no token, a cell must wait in the queue until a new token is generated. Tokens are

¹ Also referred to as policing.

generated at a fixed rate corresponding to the average rate of the connection. If the number of tokens in the token pool exceeds some predefined threshold value, the process of token generation stops. This threshold value corresponds to the burstiness of the transmission; the larger the threshold value, the bigger the burstiness. This method enforces the average input rate while allowing for a certain degree of burstiness. The Leaky Bucket method can also enforce the peak bandwidth by generating tokens at the rate corresponding to the peak rate.

In the original Leaky Bucket method proposed in [34], the input buffer is not provided. In [36], the input buffer is suggested to provide better control of the trade-off between the cell waiting times and the cell loss probabilities. In an extreme case, where no input buffer is provided, incoming cells do not have to wait in the buffer, but a large number of cells may be lost since all the violating cells are discarded. In the other extreme case (where an infinite input buffer is provided), no incoming cell will be lost, but cells may suffer a long waiting time. By choosing an appropriate input queue size, the trade-off between these two extremes can be controlled. In [37], an exact analysis of Leaky Bucket methods with and without an input queue is presented, providing the Laplace transforms for the waiting times and the inter-departure times of cells from the system (i.e., inter-departure times of tokens from a token pool). The expected waiting time, the cell loss probability, and the variance of the inter-departure times are also obtained. In this paper, a Poisson process is assumed for the cell arrival process. A Poisson process, however, may not accurately describe bursty traffic found in ATM networks.

In the Leaky Bucket method, violating cells are either discarded or stored in a buffer even when the network load is light, and thus network resources are wasted. The total network throughput can be improved by using the marking method² presented in [25, 26, 38]. In this scheme, violating cells, rather than being discarded, are permitted to enter the network with violation tags in their cell headers. These violating cells are discarded only when they arrive at a congested node. If there are no congested nodes along the routes, the violating cells are transmitted without being discarded. This marking method can easily be implemented using the Leaky Bucket method described above. When the queue length exceeds a threshold, cells are marked as “droppable” instead of being discarded. Through simulations it is shown that by choosing an appropriate threshold value, the marking method can guarantee a performance level required by non-violating cells and at the same time, can improve the network throughput. One possible disadvantage of this marking scheme is that processing time in each node is increased slightly because each node has to distinguish tagged cells from non-violating cells when the node is in a congested state. Each node must also monitor its state to determine if it is in congestion. (For instance, each node may check its queue length to detect the congested state.) However, this extra processing can be done quickly and easily, and the overall merits of the marking method far exceed its slight disadvantages.

An ideal bandwidth enforcement scheme should be able to correctly identify all the violating cells and discard or tag only violating cells. It should also be able to detect violation rapidly once it occurs. However, the bursty nature of the traffic carried in ATM

² Also referred to as a Virtual Leaky Bucket Method.

networks makes it difficult to implement such an ideal scheme. When the traffic is bursty, a large number of cells may be generated in a short period of time, yet conform to the traffic descriptor values claimed at the time of call establishment. For instance, the average cell arrival rate can be kept constant if cells do not arrive for a while, even if there is a burst of cell arrivals in a short time period. In this case, none of these cells should be considered violating cells. If a small value is used for a threshold, some of the cells will be falsely identified as violating cells; therefore, a relatively large threshold value must be used to avoid discarding or tagging non-violating cells. However, this large threshold value makes it harder to distinguish truly violating transmissions from temporary burst transmissions; thus, the time required to detect violations is increased. As a result, in an ATM environment it may be more desirable to apply a marking method in order to avoid undesired enforcement actions by the network.

Bandwidth enforcement schemes may also be used with traffic shaping³. The purpose of traffic shaping is to throttle cell inputs into a network to avoid the bursty cell transmissions. Burst cell transmissions are avoided, for example, by separating successive ATM cells by idle times. The shaping function could be performed by the access control either at a user-network interface or at a data source by buffering and injecting cells into the network at a slower speed. Since traffic shaping reduces network congestion by suppressing inputs to the network, it may be able to support a greater number of calls than a network without the shaping function. With traffic shaping, the entire transmission of traffic may be unnecessarily slowed since cells are injected into a network at a slower speed even when the network load is light. However, with traffic shaping this degradation in the service quality is achieved in a more graceful way.

2.2.3. Coping with Traffic Uncertainties

In the previous subsections, admission control and bandwidth enforcement schemes are examined. In admission control, the network performance is predicted based on the traffic descriptor values provided by the network users, and then a decision is made as to whether a new connection is accepted or not. In bandwidth enforcement, each connection is monitored, and the traffic flow is forced to conform with the traffic descriptor values provided by the network users. However, the exact traffic characteristics may not be available to the network users, and therefore, the values of traffic descriptors provided by the users may involve large uncertainty. In such a case, a network may underestimate the impact of accepting a new connection and congestion may result.

Very little attention has been paid to the problem of uncertainty in traffic descriptor values. Holtzman addressed this issue in [21], examining three approaches which were originally proposed in other contexts, and considering their application to the problem of traffic uncertainty in ATM networks. The three approaches examined by Holtzman are the approach using random variables [39], the fuzzy set approach [40] and the neural net approach to learn about the uncertain environment [41]. In this subsection, the first approach, which is the most promising and widely applicable, is discussed.

In the first approach discussed by Holtzman, the uncertainty in the traffic descriptor

³ Also referred to as traffic smoothing.

values is quantified by using a random variable for each uncertain parameter in the traffic model. Assume that the cell arrival process to the network is characterized by a point process parameterized by k traffic parameters, x_1, \dots, x_k . Further, assume that the delay incurred by cells through the network in question is a function of the k traffic descriptors and is given by $D(x_1, \dots, x_k)$. Assume that the performance requirement is given, and it is to keep the delay (mean or percentile) less than a given threshold value D^* (i.e., $D(x_1, \dots, x_k) < D^*$). Since it is assumed that the delay function $D(x_1, \dots, x_k)$ is known, we can determine the feasible parameter region Ω to satisfy the performance requirement $D(x_1, \dots, x_k) < D^*$. Ω is a range of possible values of the traffic descriptors which satisfy a given performance requirement.

Let us denote $\underline{Y}^j = (Y_1^j, \dots, Y_k^j)$ as a set of random variables which parameterize the arrival stream for the j -th network user. Using \underline{Y}^j ($j = 1, \dots, n$), the aggregated cell arrival process from n users can be obtained. Let us denote this aggregated arrival process as $\underline{X}^{(n)} = (X_1^{(n)}, \dots, X_k^{(n)})$. In general, $\underline{X}^{(n)} = f(\underline{Y}^1, \dots, \underline{Y}^n)$. $\underline{X}^{(n)}$ is a set of random variables which parameterize the cell arrival process from the superposition of n users.

From $\underline{X}^{(n)}$, the number of users n^* , which can be supported by the network satisfying the performance requirement $D(x_1, \dots, x_k) < D^*$ with high probability, can be determined. n^* is given by $n^* = \max_n \{n : P[\underline{X}^{(n)} \in \Omega] > 1 - \delta\}$, where δ is a predefined tolerance level. (For non-homogeneous superpositions, the traffic mix should be specified.)

In obtaining the aggregated arrival process $\underline{X}^{(n)}$, traffic uncertainties are considered. The process of obtaining $\underline{X}^{(n)}$ can be better illustrated using an example. Assume that the traffic generated by the j -th user is characterized by the mean cell arrival rate and the squared coefficient of variation of the time between cell arrivals. Further, assume that these two parameters have uncertainties. For the traffic generated by the j -th user, assume that:

- The mean cell arrival rate is modeled by a normally distributed random variable Y_1^j , with mean λ_j and variance $\sigma_{\lambda_j}^2$.
- The squared coefficient of variation of the time between cell arrivals is modeled by a normally distributed random variable Y_2^j , with mean c_j^2 and variance $\sigma_{c_j^2}^2$.
- The random variables Y_1^j and Y_2^j are mutually independent.

In the above, the uncertainties in the mean cell arrival rate and the squared coefficient of variation of the time between arrivals are quantified by using random variables Y_1^j and Y_2^j , respectively. Then for the superposed arrival process $\underline{X}^{(n)} = (X_1^{(n)}, X_2^{(n)})$, $X_1^{(n)}$ (the mean arrival rate) and $X_2^{(n)}$ (the squared coefficient of variation of the time between arrivals) need to be calculated. They are calculated using the QNA approximation [42]. That is,

$$X_1^{(n)} = \sum_{j=1}^n Y_1^j, \quad X_2^{(n)} = W^{(n)}(Z^{(n)} - 1) + 1, \quad (8)$$

where

$$Z^{(n)} = \frac{\sum_{j=1}^n Y_1^j Y_2^j}{\sum_{j=1}^n Y_1^j}, \quad W^{(n)} = \frac{1}{1 + 4(1 - P^{(n)})(V^{(n)} - 1)}. \quad (9)$$

$P^{(n)}$ and $V^{(n)}$ are given by

$$P^{(n)} = s \sum_{i=1}^n Y_1^i, \quad V^{(n)} = \frac{(\sum_{i=1}^n Y_1^i)^2}{\sum_{i=1}^n (Y_1^i)^2}, \quad (10)$$

where s is the mean service time.

Finally, the joint distribution of $X_1^{(n)}$ and $X_2^{(n)}$ needs to be computed. It is found that a bivariate normal distribution is a good approximation. The means and variances of random variables $X_1^{(n)}$ and $X_2^{(n)}$, and the correlation between $X_1^{(n)}$ and $X_2^{(n)}$ are approximated using a Taylor series expansion technique.

Note that although this approach allows uncertainty in parameter values, it must have *a priori* knowledge about the system model (e.g., knowledge about the arrival process and the service process).

3. Multiple Traffic Classes

In ATM networks, there is diversity of service and performance requirements. For instance, real-time voice has strict delay requirement, but loss of small amount of voice information is affordable. Many data applications, on the other hand, does not require real-time delivery, but high throughput and strict error control are required. Some video applications such as video conferencing communications require error-free transmission as well as short delays [1]. Ideally, uniform control mechanism should be applied to satisfy all the service and performance requirements. However, this is extremely difficult. As an alternative, a notion of multiple traffic classes or GOS (Grade of Service) can be introduced, and separate control mechanism can be applied depending on the traffic class. The notion of multiple traffic classes is also required since full efficiency cannot be obtained by simple statistical multiplexing when traffic load is high. Greater bandwidth efficiency can be attained through the provision of multiple traffic classes. For instance, loss-insensitive traffic may be discarded first when traffic load becomes high. The bursty nature of traffic also requires the notion of multiple traffic classes. Bursty traffic is more likely to let the network congested instantaneously. In such case, separate control mechanism is required depending on the traffic class to relieve the congestion.

To support multiple classes of traffic in ATM networks, there are two possible approaches [20]:

- **Segregation** According to the given quality class, network resources are partitioned and made available to each class independently.
- **Priority** It provides multiple priority levels depending on the class of traffic. In this approach, network resources are shared among different classes of traffic.

The priority scheme is easier to implement than segregation scheme. Furthermore, it is shown that the integrated approach performs better than the segregated approach in terms of cell loss probability and average queueing delay as far as a high priority service does not take a large portion of traffic [24]. In this section, various priority schemes are examined.

The simplest priority scheme is a fixed priority scheme. In this scheme, the priority is always given to the delay-sensitive class (i.e., real time traffic). This scheme causes relatively

high losses for the loss-sensitive traffic, while providing relatively low delays for the delay-sensitive traffic [54]. The converse holds true for FCFS, where there is no class distinction between delay-sensitive and loss-sensitive traffic, and cells are served in the order of arrival. Various dynamic priority disciplines have been proposed. In [54], the performance of two dynamic priority disciplines - Minimum Laxity Threshold (MLT) and Queue Length Threshold (QLT) - in a statistical multiplexer is examined. In Minimum Laxity Threshold (MLT) discipline, the priority is given to delay-sensitive traffic when the minimum laxity (defined as the amount of time until the first deadline of all queued real-time packets expires) among the queued delay-sensitive cells is less than or equal to some threshold value; otherwise the priority is given to the loss-sensitive traffic. The laxity is defined as the remaining number of slots until its deadline expires. A packet remains in the queue until either the packet is transmitted or the laxity reaches zero. When the laxity reaches zero, the packet is considered as lost. In Queue Length Threshold (QLT) discipline, priority is given to loss-sensitive traffic when the number of loss-sensitive cells in the queue exceed some threshold value; otherwise the priority is given to the delay-sensitive traffic. Here, the analytical and numerical performance models for the above disciplines are developed, and the following observations are made:

- Both the MLT and QLT policies allow the designer to explicitly tradeoff the performance realized by each class, choosing an appropriate value for the threshold.
- Little difference in the performance tradeoffs is observed in the MLT policy and the QLT policy. Due to the simpler implementation, it is concluded that the QLT policy is more practical than the MLT policy.

In [55], a dynamic priority discipline, Head-of-the-Line with Priority Jumps, for a multiplexer with multiple classes of delay-sensitive traffic is studied. Here, it is assumed that each priority class has its own separate queue. Within each queue, cells are FCFS, while high-priority queue has nonpreemptive priority over lower-priority queues. However, a limit is imposed on the maximum queueing delay of cells at each queue. When the maximum delay limit is exceeded, that cell jumps to the end of the next higher priority queue. Thus, the queueing delay of a cell before it joins the highest priority queue is bounded by the summation of the delay limits at all the queues with priorities equal to or higher than the cell's original class. The performance of the different classes can be controlled by adjusting the values of these delay limits. It is shown that this discipline minimizes the overall maximum tail probability in the system, and thus, it is suitable for a network carrying voice and video traffic, for which the tail behavior of the delay distribution has significant effect on the performance.

In [7], two dynamic priority disciplines - Oldest-Customer-First (OCF) and Earliest-Deadline-First (EDF) - are proposed to minimize the variability of cell delays in voice traffic. Under the Oldest-Customer-First (OCF) discipline, the priority of a cell depends on its age (the time it has spent so far in the network). On an arrival of a cell to a queue, its age is determined from its timestamp and compared to the ages of cells already in the queue. The arriving cell is inserted behind older cells and ahead of younger cells. Under the Earliest-Deadline-First (EDF) discipline, the priority of a cell depends on its "deadline", determined as $t + (D_{max} - a - d)$, where t is arrival time, D_{max} is maximal allowable entry-to-exit delay,

a is the cell age, and d is the estimated delay along its remaining route in the network. d depends on route length and the traffic conditions at other queues. Similar discipline was already described by Jackson [56-58]. In Jackson's Earliest-Due-Date (EDD) (or Dynamic) discipline, the priority of a customer depends on their due-dates, defined as the sum of his arrival time to the system and his "urgency number". The server chooses a customer whose due-date is closest to the current time, or if there is any customer with "over-due", the server chooses the one with the largest "over-due". It is different from the EDF discipline since in EDD, priority is a function of only waiting time and not of other factors such as route or traffic conditions. The EDF discipline can be applied to integrated packet networks where voice and data traffic are carried. In that case, since voice traffic has stricter expiration times than data traffic, voice traffic will have stricter deadlines and thus have higher priority under EDF. The authors, however, point out possible practical disadvantages of these disciplines - the expensive processing and sorting necessary at each queue.

Priority schemes can also be used as a congestion control protocol. With no priority scheme, arriving cells are discarded without bias when the queue is full. This method, however, fails to recognize different requirements of different classes of traffic. For instance, data traffic must generally be received error-free, whereas the inherent structure of speech allows for some loss of information without significant quality degradation. In [19], network provides two different services - express service and first-class service. Express service is a service which requires real-time delivery, but can tolerate some loss of information. It is appropriate for delay-sensitive class such as voice and some video traffic. First-class service is a service which does not require real-time delivery, but requires high throughput and strict error control. It is appropriate for loss-sensitive class such as data. In case of congestion, higher priority is given to delay-sensitive class (express-class) traffic, and reactive control is performed only on loss-sensitive class (first-class) traffic in the form of choking/relieving. To overcome the problem of reactive control, that is, to overcome the problem caused by slow feedback, the reactive control is exercised only at the edge of the network, and no control is performed at intermediate nodes. No congestion control is performed on higher-priority class (i.e., delay-sensitive class). Here, two separate buffers are used for each class of traffic, and congestion is detected if the queue length of loss-sensitive class exceeds some threshold value. This mechanism of detecting congested state may not be accurate in an ATM environment where a number of bursty traffics change the queue length very rapidly.

In [63], two priority schemes - variable priority method A and variable priority method B - are examined to realize delay, or loss sensitive service classes. Here, it is assumed that each output line of the switch has a buffer associated with it, and common buffer is used for the two service classes. In both methods, the loss-sensitive class always has priority regarding cell input with buffer overflow; when the buffer is overflowed, delay-sensitive class cells within the buffer are discarded when required for the loss sensitive class input. Two priority methods differ only regarding cell output from buffer to line. In variable priority method A, the delay-sensitive class always has priority regarding cell output from buffer to line, whereas in variable priority method B, output priority is changed from the delay sensitive class to the loss sensitive class when loss sensitive class cells wait for longer than a determined threshold period within the buffer. Therefore, this method improves the both delay and loss performance of the loss-sensitive class. This method is similar to the QLT

discipline discussed above, but in the QLT policy, the threshold parameter is the queue occupancy of loss-sensitive class cells, whereas in the variable priority scheme B, it is the time spent in the queue.

Within the same real-time stream, some data may have greater value than other data [9]. For example, in coded speech, active speech usually carries more important informations than background noise during pauses does. In [8-11], congestion controls based on priority-oriented packet discarding of speech are examined in integrated packet networks. They are all based on selectively discarding voice packets whose loss will have the least effect in quality of the reconstructed voice signal. Priority is assigned to each packet at the transmitter, and low priority packets are dropped when congestion happens. It is shown that such a prioritized system is capable of achieving better performance than non-prioritized systems with the same delay time and packet loss constraints [11]. In [8], impact of voice packet discarding on data traffic is studied. It is shown that the mean waiting time for data can be significantly reduced by selectively discarding voice packets during periods of congestion.

There are three methods of determining packet priorities. They are:

- **embedded coding** [12] In this method, the encoded information is divided into more significant bits and less significant bits. Separate packets are formed according to their importance. Priority is given to packets containing more significant bits.
- **even/odd samples** [14] In this method, speech samples are identified as either even or odd. The even samples form high priority packets and the odd form low priority packets (or vice versa).
- **multiple energy detection thresholds** In this method, packets are classified as "semi-silence" and "active" according to their energy level. The energy level is estimated and compared to two thresholds. If it lies below both thresholds, it is assumed to be a silence, and no packet is formed. If it lies between two thresholds, "semi-silent" packets are formed. If it lies above both thresholds, "active" packets are formed. Priority is placed on "active" packets.

Two types of thresholds are used as a measure of overload [11]. They are:

- **speaker activity threshold** Low-priority packets are discarded when the number of callers in talkspurt exceeds a threshold.
- **buffer content threshold** Low-priority packets are discarded when the queue length in the multiplexer exceeds a threshold.

The effect of particular discarding algorithms and associated overload measures on network performance needs to be studied further.

Slightly different techniques to control voice traffic have been proposed. In [15 - 18], priority is assigned to more important (significant) bits, not to cells. Each cell consists of high priority bits (more significant bits) and low priority bits (less significant bits), and cell size is reduced in response to overload by dropping low priority bits. This technique has a major disadvantage: it requires network nodes to know the internal structure of a voice cell in order to distinguish high priority bits from low priority bits and to manipulate the cell contents [9]. This will increase cell processing at each switching node; thus, this technique

may not be suitable for ATM networks. Furthermore, since the cell size is constant in ATM networks, it is not clear how this technique can be applied in ATM networks.

Discarding cells based on the importance of their contents can also be applied to video traffic. If an embedded coding technique⁴ [59-61] is used for the image, coded information is separated into two bit streams: a stream containing essential information and a stream containing picture enhancement information. Cells containing essential information are given higher priority than those containing the picture enhancements. When congestion occurs, only low priority cells are discarded. With this scheme, even when networks become congested the essential parts of coded information are transmitted; thus, it is expected that cell loss will have only a small influence on picture quality [60].

In this section, we have examined various priority schemes to provide multiple classes of traffic. In an integrated network, the followings are the possible traffic classes subject to priority schemes:

- video,
- voice,
- embedded coded video or voice,
- data, and
- tagged cells which violate traffic descriptions

How to put priorities on traffic classes in a network where all the above traffic classes are present is a possible research topic.

Finally, it should be noted that priority schemes may not be appropriate in an integrated network where a large portion of traffic is a high priority service [64]. In such case, low priority traffic will have severe performance degradation. In this situation, a segregation scheme (i.e., physical traffic separation) may work better. Performance comparison between these two schemes in such environment needs to be studied further.

4. Error Control

Due to the use of optical fibers, the ratio of propagation delay to cell transmission time and the ratio of processing time to cell transmission time have increased in ATM networks. The use of optical fibers has also resulted in significant reduction in the channel error rate. These changes make it necessary to re-examine the error control schemes used in existing networks.

Before examining error control schemes, it should be noted that flow control (which is conventionally coupled with error control as a form of window flow control) must be independent of error control in high speed networks such as ATM networks [43]. The use of windows for both flow control and error control leads to a conflict. In a high speed network, where the ratio of propagation delay to cell transmission time is large, a large window must be used to achieve high throughput; however, a large window imposes little control effect. In the worst case, an entire window may be transmitted at once, possibly leading to network

⁴ Also referred to as a layered coding technique or a hierarchical coding technique.

congestion. Therefore, some form of rate-based flow control schemes such as the Leaky Bucket method discussed in subsection 2.2.2 must be used independently of error control as a flow control scheme. In this section, various error control schemes are examined under a high speed environment.

Error control schemes can be implemented on a link-by-link basis or an edge-to-edge basis. In a link-by-link scheme, retransmission of lost or erred cells takes place only between adjacent switching nodes, whereas in an edge-to-edge scheme, retransmission takes place only between the source and destination nodes. The suitability of link-by-link and edge-to-edge schemes in high speed networks is discussed in subsection 4.1.

Error control schemes can be placed into two classes depending on the retransmission protocols (Automatic Repeat Request (ARQ) protocols) for erred/lost cells: go-back-n and selective-repeat protocols. In both go-back-n and selective-repeat protocols, the transmitter sends cells continuously without waiting for an acknowledgement from the receiver. Upon receipt of a negative acknowledgement (NAK) or when timeout occurs, a go-back-n protocol retransmits all the cells starting with the lost cell, whereas a selective-repeat protocol retransmits only the lost cell. In subsection 4.2, these schemes are further discussed with possible improvements under a high speed environment.

4.1 Link-by-Link vs. Edge-to-Edge Schemes

In traditional packet networks (e.g., X.25/X.75), error control is done on a link-by-link basis. Link-by-link error control, however, may not be appropriate in high speed networks such as an ATM network. Link-by-link schemes involve heavy protocol processing because cells are manipulated and processed at each node in the network. The overhead of protocol processing is very significant in high speed networks, while in existing networks, this overhead is considered to be negligible (a packet length and a channel speed are the bottleneck in determining the packet transmission delay). Therefore, in high speed networks, edge-to-edge schemes may become more attractive, despite the fact that it “wastes” the successful transmissions over all earlier links if error or loss happens later on the path [44]. In edge-to-edge schemes, an erred or lost cell is retransmitted from the source, and thus, if errors/losses occur on a link far from the source, edge-to-edge retransmission wastes the successful transmissions over all earlier links.

Some papers investigate the issue of link-by-link versus edge-to-edge schemes [44 - 47]. In [44], the performance of these two schemes are investigated and compared. The network is modeled as a tandem queueing network, where each queue represents a single switching node along a virtual circuit. Finite buffers are assumed except at the source node (infinite buffers are assumed at the source). Blocked cells due to buffer overflow are considered lost and have to be retransmitted. The effects of propagation delay are considered, while the processing times required in an error recovery protocol are assumed to be negligible. From a mathematical analysis and simulations, they concluded the following:

- For small virtual circuit (VC) throughput values, there is no significant performance difference between the link-by-link and the edge-to-edge schemes.
- As the VC throughput increases, the edge-to-edge scheme performs better than the link-by-link scheme. The edge-to-edge scheme experiences smaller delays and reaches the

saturation point at a higher throughput value than the link-by-link scheme. This is because the edge-to-edge scheme buffers a copy of a cell only at the source and releases the buffer immediately upon completion of cell transmission at the intermediate nodes, whereas the link-by-link scheme buffers each cell until it is acknowledged. Therefore, for high throughput values, blocking due to buffer overflow happens less frequently in the edge-to-edge scheme, yielding lower delays and a higher maximal throughput.

- When the error probability is increased, the link-by-link scheme performs marginally better than the edge-to-edge scheme at lower throughput values. This is because when an error occurs, link-by-link recovers faster than edge-to-edge since the link-by-link scheme receives feedback earlier than the edge-to-edge scheme. However, as the throughput increases, blocking due to buffer overflow becomes significant, and thus, the edge-to-edge scheme eventually achieves better performance.
- When the propagation delay is decreased, the relative performance of a link-by-link scheme improves. (The opposite holds when the propagation delay is increased, or equivalently, when the channel speed is increased.) This is because when the propagation delay is decreased, the intermediate nodes need to buffer a copy of a cell for a shorter time, and thus the blocking probabilities for the link-by-link case decrease.
- As the number of hops increases, the relative difference in the performance of the two schemes decreases. This is due to the disadvantage of edge-to-edge schemes discussed earlier: if errors/losses occur on a link far from the source, edge-to-edge schemes waste the successful transmissions over all earlier links. It is conjectured that, as the number of hops increases, the performance of link-by-link schemes will eventually become better than that of edge-to-edge, but this crossover occurs when the number of hops becomes unrealistically large.

In summary, the authors conclude that even under assumptions that favor the link-by-link scheme (e.g., no node processing time is considered, and an analytic model which overestimates the delay for the edge-to-edge schemes is used), the edge-to-edge scheme performs better than the link-by-link scheme, requiring fewer network resources such as buffers and computation time.

The performance of link-by-link and edge-to-edge schemes in a high speed environment is also investigated in [45]. Here, the effects of processing time required for error recovery are considered, while the effects of propagation delay are assumed to be negligible. The network is modeled as a tandem queueing network with feedback loops between adjacent nodes (for a link-by-link scheme), and the source and destination nodes (for an edge-to-edge scheme). Each queue represents a protocol layer within a switching node, rather than a switching node as a whole. Infinite buffers are assumed at each switching node. It is concluded that for a network with very high-speed and low-error-rate channels, the edge-to-edge scheme gives better performance (i.e., gives the smaller cell transmission delay and cell loss probability) than the link-by-link scheme. The analytic models used in this paper are validated in [46] through simulations.

In [47], link-by-link and edge-to-edge schemes are studied through simulations for existing X.25 packet networks. The propagation delays and processing times are considered, although they are relatively small (on the same order as a packet transmission time). It is concluded

that the edge-to-edge protocol alone can yield good performance if the value for the edge-to-edge timeout is large enough to avoid unnecessary timeouts (i.e., timeouts under heavy loads on error-free transmissions). It is also shown that the addition of a link-by-link control significantly improves the network performance under light loads. The improvement gained by adding link-by-link control is reduced as the load increases. This is because under a light load, the link-by-link control causes little queueing since there is not much traffic, and thus, very few edge-to-edge timeouts occur. On the contrary, when traffic load is heavy, the additional link processing causes queues to build up, which in turn causes edge-to-edge timeouts.

In conclusion, in ATM networks, where the effects of propagation delay and processing time are significant, the edge-to-edge scheme gives better performance than the link-by-link scheme.

4.2. Go-Back-N vs. Selective-Repeat Protocols

The channel propagation delay has a strong effect on the performance of ARQ protocols. As the ratio of propagation delay to cell transmission time increases, the go-back-n protocol suffers from the reduced throughput because of the large number of cell retransmissions required. The selective-repeat protocol achieves better throughputs than the go-back-n protocol since it retransmits only those cells that are negatively acknowledged or whose timeouts have expired. However, in the selective-repeat protocol, a reordering buffer is required at the receiver since cells need to be buffered until all preceding cells are received correctly.

In existing networks such as X.25 networks, the go-back-n protocol is the most commonly used protocol since it eliminates the need for reordering buffers and is easy to implement. Furthermore, this protocol provides reasonably good performance when the propagation delay is comparable to the packet transmission time [48]. However, in ATM networks where the channel speed is very high and the ratio of propagation delay to cell transmission time is very large, the number of cells in transit can be very large, and the go-back-n protocol may not perform well. For example, as we saw in the example given at the beginning of section 3, if we use 500 bit long cells, 5 μ sec/Km propagation delay and 1 Gbits/sec channel, and if the transmitter and the receiver are 100 Km apart, there are 1,000 cells in transit from the transmitter to the receiver. Therefore, if a go-back-n protocol is used, upon receipt of a NAK, the transmitter may need to retransmit 2,000 cells. This can cause significant reduction in throughput. Thus, the go-back-n protocol becomes less attractive, and the selective-repeat becomes more attractive in a high speed network environment.

In [48], a throughput analysis of a selective repeat protocol in high-speed network environments is presented. Through numerical examples it is shown that as the ratio of propagation delay to cell transmission time increases, the selective-repeat achieves significantly better throughput performance than a go-back-n protocol. However, at the same time, the buffers required for reordering cells become large. This effect is studied in [49], and the upper bound on the mean buffer occupancy is derived.

The trade-off between the throughput and the buffer requirement should be carefully weighed to determine which ARQ protocol is more effective in a high speed environment. However, with the advancement in VLSI technology, cost-effective reordering buffers are in fact coming to the marketplace [50], and the implementation of selective-repeat is believed

to be simple as long as the cell size is constant [51]. Therefore, the current trend is to favor the selective-repeat protocol over the go-back-n protocol.

In the above, the effects of propagation delay on the performance of ARQ protocols are examined. Another important factor which needs to be addressed is the use of optical fiber and its extremely low error rate. The effects of low error rate on the performance of ARQ protocols also need to be considered. As the error rate decreases, errors rarely happen, and thus, the inefficiency of the go-back-n protocol becomes less significant since retransmissions are rarely required. Even though the selective-repeat protocol always achieves better throughput, the performance difference between the selective-repeat and go-back-n protocols becomes smaller as the error rate on the channel decreases. The effects of large propagation delay and low error rate need to be carefully examined to determine the best ARQ scheme for ATM networks.

Finally, it should be noted that to improve selective-repeat protocols a "block" concept can be introduced. A block is a group of cells, and a single acknowledgement message is used to acknowledge a block of cells, not an individual cell. Selective-repeat is performed on a block basis, and either the entire block can be retransmitted or, if an acknowledgement contains the list of cells to be retransmitted, only the erred/lost cells can be retransmitted.

Several forms of block acknowledgement schemes have been proposed in [43, 51 - 53]. All the block acknowledgement schemes share a common advantage over a per-cell acknowledgement scheme: they reduce the large overhead incurred by sending a separate acknowledgement for each cell. Furthermore, the number of bits required to address the erred/lost cells is reduced if the entire block is retransmitted, even when one cell in a block is delivered incorrectly [53]. This reduction in the address field results in a reduction in processing time by reducing the size of tables which have to be searched to determine which blocks have to be retransmitted. This effect of reduced processing time is more significant in high speed networks since minimum processing is required to match the rapid channel speed.

In conclusion, the authors of this paper believe that a block acknowledgement scheme in conjunction with a block-based selective-repeat retransmission protocol executed on an edge-to-edge basis is the most appropriate error control scheme in very high speed networks such as ATM.

5. Conclusion and Future Research Problems

Among the techniques proposed for B-ISDN transfer mode, ATM concept is considered to be the most promising transfer technique because of its flexibility and efficiency. In this paper, a number of topics related to ATM networks are surveyed and reviewed. Those topics covered in this paper are congestion control, provision of multiple classes of traffic, and error control. The following summarizes the contents of this paper:

- Due to the effects of high speed channels, conventional congestion schemes based on feedback information may not work in ATM networks. For the same reason, preventive control is more effective in ATM networks than reactive control.
- Due to the diversity of service and performance requirements, the notion of multiple traffic classes is required, and separate control mechanism should be used depending on the traffic classes. The most effective method to support multiple classes of traffic is the

priority scheme.

- Fundamental changes due to the use of optical fibers trigger the necessity to reexamine the conventionally used error control schemes. A block acknowledgement protocol in conjunction with a selective-repeat retransmission scheme executed in edge-to-edge basis is the most promising candidate for the error control scheme.

Finally, it should be noted that there are still a number of unsolved problems in ATM network. The following presents some of the future research problems:

- In most of the past analytic work, homogeneous traffic sources are assumed. Quantitative performance analysis assuming heterogeneous traffic sources needs to be studied further.
- In the past analysis of leaky bucket methods, Poisson arrivals are assumed. This is not a valid assumption in ATM networks. An analysis of leaky bucket methods considering the burstiness of traffic needs to be done.
- A Virtual Leaky Bucket Method (i.e., a leaky bucket with marking method) needs to be analyzed to investigate its performance.
- In most of the work, congestion is detected if the queue length exceeds some threshold value. This mechanism of detecting congested state may not be accurate in an ATM environment where a number of bursty traffics change the queue length very rapidly. Thus, more efficient ways to detect congestion need to be investigated.
- So far, cell loss probability has been investigated on the multiplexed link. Cell loss probability on each connection not on the multiplexed link needs to be further investigated to find out the performance seen by end users.
- Performance comparison between priority and segregation schemes in an integrated network where a large portion of traffic is a high priority service needs to be done.
- How to put priorities on traffic classes in a network where all possible traffic classes are mixed is also a possible research topic.

References

- [1] H. Ichikawa, M. Aoki and T. Uchiyama, "High-Speed Packet Switching Systems for Multimedia Communications," IEEE J. Select. Areas in Commun., Vol.SAC-5, No.8, pp.1336-1345, October 1987.
- [2] A. Hac and H. B. Mutlu, "Synchronous Optical Network and Broadband ISDN Protocols," Computer, Vol.22, No.11, pp.26-34, November 1989.
- [3] R. Handel, "Evolution of ISDN Towards Broadband ISDN," IEEE Network, pp.7-13, January 1989.
- [4] T. Kamitake and T. Suda, "Evaluation of an Admission Control Scheme for an ATM Network Considering Fluctuations in Cell Loss Rate," Proc. IEEE GLOBECOM '89, pp.49.4.1-49.4.7.
- [5] M. Murata, Y. Oie, T. Suda and H. Miyahara, "Analysis of a Discrete-Time Single-Server Queue with Bursty Inputs for Traffic Control in ATM Networks," Proc. IEEE GLOBECOM '89, pp.49.5.1-49.5.7.
- [6] S-Q. Li, "Study of Information Loss in Packet Voice Systems," IEEE Trans. Commun., Vol.37, No.11, pp.1192-1202, November 1989.
- [7] T. M. Chen, J. Walrand and D. G. Messerschmitt, "Dynamic Priority Protocols for Packet Voice," IEEE J. Select. Areas Commun., Vol.7, No.5, June 1989.
- [8] N. Yin, S-Q. Li and T. E. Stern, "Data Performance in an Integrated Packet Voice/Data System Using Voice Congestion Control," Proc. IEEE GLOBECOM '88, pp.16.4.1-16.4.5.
- [9] D. W. Petr, L. A. DaSilva, Jr. and V. S. Frost, "Priority Discarding of Speech in Integrated Packet Networks," IEEE J. Select. Areas Commun., Vol.7, No.5, pp.644-659, June 1989.
- [10] N. Yin, T. E. Stern and S-Q. Li, "Performance Analysis of a Priority-Oriented Packet Voice System," Proc. IEEE INFOCOM '87, pp.856-863.
- [11] N. Yin, S-Q. Li and T. E. Stern, "Congestion Control for Packet Voice by Selective Packet Discarding," Proc. IEEE GLOBECOM '87, pp.45.3.1-45.3.4.
- [12] D. J. Goodman, "Embedded DPCM for variable bit rate transmission," IEEE Trans. Commun., Vol. COM-28, pp.1040-1046, July 1980.
- [13] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson and J. D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communications," IEEE Trans. Commun., Vol.36, No.7, pp.834-844, July 1988.
- [14] N. S. Jayant and S. W. Christensen, "Effects of packet losses in waveform coded speech and improvements due to an odd-even sample-interpolation procedure," IEEE Trans. Commun., Vol. COM-29, pp.101-109, February 1981.
- [15] K. Sriram and D. M. Lucantoni, "Traffic Smoothing Effects of Bit Dropping In a Packet Voice Multiplexer," Proc. IEEE INFOCOM '88, pp.8A.1.1-8A.1.12.

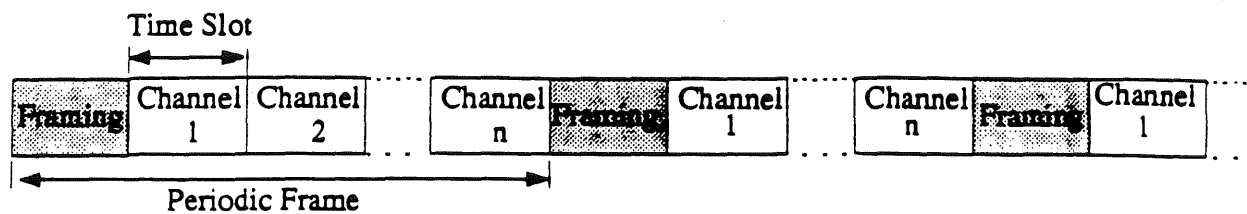
- [16] V. R. karanam, K. Sriram and D. O. Bowker, "Performance Evaluation of Variable-Bit-rate Voice in Packet-Switched Networks," AT&T Tech. J., pp.57-69, September/October, 1988.
- [17] S. Dravida and K. Sriram, "End-to-End Performance Models for Variable Bit Rate Voice over Tandem Links in Packet Networks," Proc. IEEE INFOCOM '89, pp.1089-1095.
- [18] R. W. Muise, T. J. Schoenfeld and G. H. Zimmerman, "Experiments with Wideband Packet Technology," Proc. 1986 Intl. Zurich Seminar Digital Commun., paper D4, March 1986.
- [19] A. Gersht and K. J. Lee, "A Congestion Control Framework for ATM Networks," Proc. IEEE INFOCOM '89, pp.701-710.
- [20] G. M. Woodruff, R. G. Rogers and P. S. Richards, "A Congestion Control Framework for High-Speed Intergrated Packetized Transport," Proc. IEEE GLOBECOM '88, pp.7.1.1-7.1.5.
- [21] J. M. Holtzman, "Coping with Broadband Traffic Uncertainties: Statistical Uncertainty, Fuzziness, Neural Networks," presented at IEEE Workshop on Computer Commun., Data Point, CA, October 1989.
- [22] L. Dittmann and S. B. Jacobsen, "Statistical Multiplexing of Identical Bursty Sources in an ATM Network," Proc. IEEE GLOBECOM '88, pp.39.6.1-39.6.5.
- [23] J. Kulzer and W. Montgomery, "Statistical Switching Architecture for Future Services," Proc. ISS '84, pp.43A.1.1-43A.1.6.
- [24] T. Y. Choi, "Statistical Multiplexing of Bursty Sources in an ATM Network," Multimedia '89.
- [25] G. Gallassi, G. Rigolio and L. Fratta, "ATM: Bandwidth Assignment and Bandwidth Enforcement Policies," Proc. IEEE GLOBECOM '89, pp.49.6.1-49.6.6.
- [26] M. Hirano and N. Watanabe, "Characteristics of a Cell Multiplexer for Bursty ATM Traffic," Proc. IEEE ICC '89, pp.13.2.1-13.2.5.
- [27] S. Akhtar, "Congestion Control in a Fast Packet Switching Network," Master's Thesis, Washington University, December 1987.
- [28] J. Y. Hui and E. Arthurs, "A Broadband Packet Switch for Integrated Transport," IEEE J. Select. Areas Commun., SAC-5, No.8, pp.1264-1273, October 1987.
- [29] K. Sriram, W. Whitt, "Characterizing Superposition Arrival Processes in packet Multiplexers for Voice and Data," IEEE J. Select. Areas Commun., SAC-4, No.6, pp.833-846, September 1986.
- [30] A. E. Eckberg, "Generalized Peakedness of Teletraffic Processes," Proc. 10th Intl. Teletraffic Congress, Montreal, 1983.
- [31] S. B. Jacobsen, K. Moth, L. Dittmann and K. Sallberg, "Load Control in ATM Networks," submitted to ISS '90.
- [32] K. Nakamaki, M. Kawakatsu and A. Notoya, "Traffic Control for ATM Networks," Proc.

IEEE ICC '89, pp.22.5.1-22.5.5.

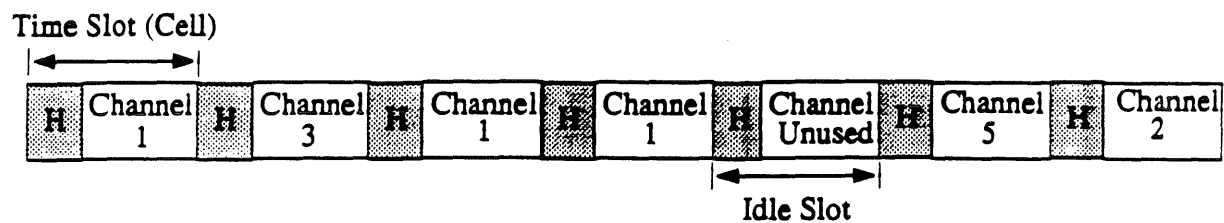
- [33] H. Esaki, Y. Katsube, K. Iwamura and T. Kodama, "A Study on Connection Admission Control for an ATM Network," Technical Report, R & D Center, Toshiba Corp., 1989.
- [34] J. S. Turner, "New Directions in Communications (or Which way to the Information Age?)," IEEE Commun. Magazine, Vol.25, No.10, pp.8-15, October 1986.
- [35] M. W. Beckner, T. T. Lee and S. E. Minzer, "A Protocol and Prototype for Broadband subscriber access to ISDNs," Proc. ISS '87, pp.B6.3.1-B6.3.8.
- [36] I. Cidon and I. S. Gopal, "PARIS: An Approach to Integrated High-Speed Private Networks," Int. J. Digital & Analog Cabled Systems, Vol.1, No.2, pp.77-86, April-June, 1988.
- [37] M. Sidi, W. Z. Liu, I. Cidon and I. Gopal, "Congestion Control Through Input Rate Regulation," Proc. IEEE GLOBECOM '89, pp.49.2.1-49.2.5.
- [38] A. E. Eckberg, Jr., D. T. Luan and D. M. Lucantoni, "Meeting the Challenge: Congestion and Flow Control Strategies for Broadband Information Transport," Proc. IEEE GLOBECOM '89, pp.49.3.1-49.3.5.
- [39] K. S. Meier-Hellstern and P. E. Wirth, "Packet Switch Provisioning Accounting for uncertainty in the Forecasting of Customer Characteristics," Proc. 7th Intl. Teletraffic Congress Specialists Seminar, Adelaide, Australia, September 1989.
- [40] R-J. Li and E. S. Lee, "Analysis of a Fuzzy Queue," Computers math & Applic., Vol.17, No.7, pp.1143-1147, 1989.
- [41] A. Hiramatsu, "ATM Communications Network Control by Neural Network," Intl. Joint Conf. on Neural Networks '89, Wash. DC, June 1989.
- [42] W. Whitt, "The Queueing Network Analyzer," Bell Syst. Tech. J., Part 1, Vol.62, No.9, pp.2779-2815, 1983.
- [43] D. D. Clark, M. L. Lambert and L. Zhang, "NETBLT: A High Throughput Transport Protocol," Proc. ACM SIGCOMM '87, pp.353-359.
- [44] A. Bhargava, J. F. Kurose, D. Towsley and G. Vanleemput, "Performance Comparison of Error Control Schemes in High-Speed Computer Communication Networks," IEEE J. Select. Areas Commun., Vol.6, No.9, pp.1565-1575, December 1988.
- [45] T. Suda and N. Watanabe, "Evaluation of Error Recovery Schemes for a High-Speed Packet Switched Network: Link-by-Link Versus Edge-to-Edge Schemes," Proc. IEEE INFOCOM '88, pp.722-731.
- [46] T. Bradley and T. Suda, "Performance of Error Recovery Schemes in a Fast Packet Switching Network," to appear in Proc. IEEE ICC '90.
- [47] P. Brady, "Performance of an Edge-to-Edge Protocol in a Simulated X.25/X.75 Packet Network," IEEE J. Select. Areas Commun., Vol.6, pp.190-196, January 1988.
- [48] H. Ahmadi, P. Kermani and P. Tran-gia, "Throughput Analysis of a Class of Selective Repeat Protocols in High-Speed Environments," Proc. IEEE GLOBECOM '89,

pp.26.4.1-26.4.9.

- [49] N. Shacham, "Queueing Analysis of a Selective-Repeat ARQ Receiver," Proc. IEEE INFOCOM '87, pp.512-520.
- [50] M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis*, Addison-Wesley Publishing Company, pp.124, 1987.
- [51] W. Zwaenepoel, "Protocols for Large Data Transfers over Local Networks," Proc. ACM SIGCOMM '85, pp.22-32.
- [52] G. M. Brown, M. G. Gouda and R. E. Miller, "Block Acknowledgement: Redesigning the Window Protocol," Proc. ACM SIGCOMM '89, pp.128-135.
- [53] K. Sabnani and A. Netravali, "A High Speed transport Protocol for Datagram/Virtual Circuit Networks," Proc. ACM SIGCOMM '89, pp.146-157.
- [54] R. Chipalkatti, J. F. Kurose and D. Towsley, "Scheduling Policies for Real-Time and Non-Real-Time Traffic in a Statistical Multiplexer," Proc. IEEE INFOCOM '89, pp.774-783.
- [55] Y. Lim and J. Kobza, "Analysis of a Delay-Dependent Priority Discipline in a Multi-Class Traffic Packet Switching Node," Proc. IEEE INFOCOM '88, pp.9A.4.1-9A.4.1.10.
- [56] J. R. Jackson, "Some Problems in Queueing with Dynamic Priorities," Naval Res. Log. Quart., Vol.7, No.3, 1960.
- [57] J. R. Jackson, "Queues with Dynamic Priority Discipline," Management Science, Vo.8, No.1, 1961.
- [58] J. R. Jackson, "Waiting Time Distributions for Queues with Dynamic Priorities," Naval Res. Log. Quart., Vo.9. No.1, 1962.
- [59] W. Verbiest, L. Pinnoo and B. Voeten, "The Impact of the ATM Concept on Video Coding," IEEE J. Select. Areas Commun., Vol.6, No.9, pp.1623-1632, December 1988.
- [60] F. Kishino, K. Manabe, Y. Hayashih and H. Yasuda, "Variable Bit-Rate Coding of Video Signals for ATM Networks," IEEE J. Select. Areas Commun., Vol.7, No.5, pp.801-806, June 1989.
- [61] M. Ghanbari, "Two-Layer Coding of Video Signals for VBR Networks," IEEE J. Select. Areas Commun., Vol.7, No.5, pp.771-781, June 1989.
- [62] S. E. Minzer, "Broadband ISDN and Asynchronous Transfer Mode (ATM)," IEEE Commun. Magazine, Vol.27, No.9, pp.17-24, September 1989.
- [63] H. Ohnishi, T. Okada and K. Noguchi, "Flow Control Schemes and Delay/Loss Tradeoff in ATM Networks," 1988 IEEE.
- [64] M. R. Wernik, "Architecture and Technology Considerations for Multimedia Broadband Communications," Proc. of Globecom'88, pp.21.2.1-21.2.5.



(a) STM Multiplexing



(b) ATM Multiplexing



Figure 1 STM and ATM Principles

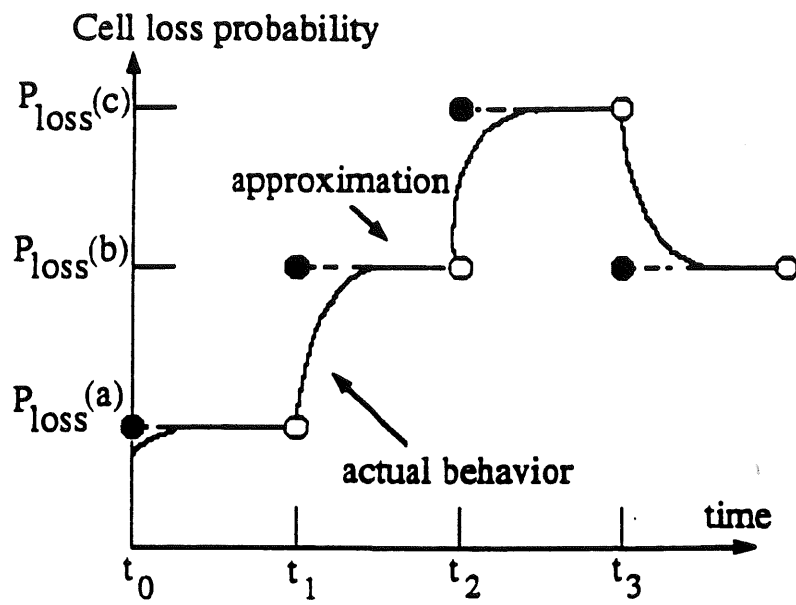


Figure 2 Time Dependent Behavior of Cell Loss Probability

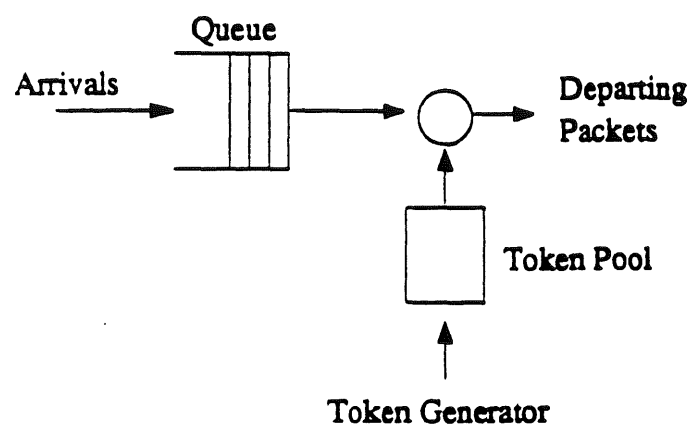


Figure 3 A Queueing Model for a Leaky Bucket Method



3 1970 00882 4739